ANNALI DI GEOFISICA, VOL. XXXVII, N. 5, September 1994

Automatic processing of seismic events recorded on a mini-array
Signal analysis combined with neural networks

Alexandre Bottero, Yves Cansi and Bernard Massinon
Laboratoire de Détection et de Géophysique, Bruyeres-Le-Chatel, France

Abstract
We present a new method for automatic processing of mini-array records of regional events. It is based on a comprehensive analysis of the cross-correlation functions. This leads to a set of time-delays used to compute the azimuth and velocity of the travelling wave only in case of consistency of the time-delay set. The second step takes into account the time-frequency representations of these wave parameters to identify each regional wave using a neural network. The resulting standard error on azimuth is 3° and the relative error on distance is less than 20%.

Key words *neural network – mini-array – automatic seismic processing*

1. Method

Among the different methods used to automatically locate earthquakes recorded on a mini-array, the method based on the broad band f-k analysis is the most common (*e.g.*: Mykkeltveit *et al.*, 1990). However, this method assumes that the propagation of the considered wave front can be modeled by a plane-wave at the scale of the array. Furthermore, in the case of local model variations (*e.g.*: station anomalies) the resulting precision is very low because of the large wavelength compared to the array extension.

To take into account these difficulties, we tested another method based on a high precision determination of the arrival-times, from which the event location is derived by a classical Husebye's method (Husebye, 1969) which computes the velocity and the azimuth of the wave. Each time-window is processed by this method in different frequency bands and activated only if the arrival-time delay set is consistent.

Furthermore, this consistency is used as a signal detector and leads to three time-frequency functions, representing first the consistency and second the velocity and azimuth when they are available. In the case of teleseismic events, the event location is strictly derived from the determinations of the velocity and the azimuth. But for regional events a phase identification is needed. This task is performed by a neural network which uses the three time-frequency functions as inputs and which leads to an estimate of each phase occurrence possibility as a function of time.

2. Data analysis and processing

In 1992, the French Laboratoire de Détection et de Géophysique of the Commissariat à l'Energie Atomique installed in Central France a temporary mini-array composed of 5 vertical short-period seismometers (fig. 1). Numerical signals are digitized at a rate of 50 samples per second with a 12-bit dynamic range.

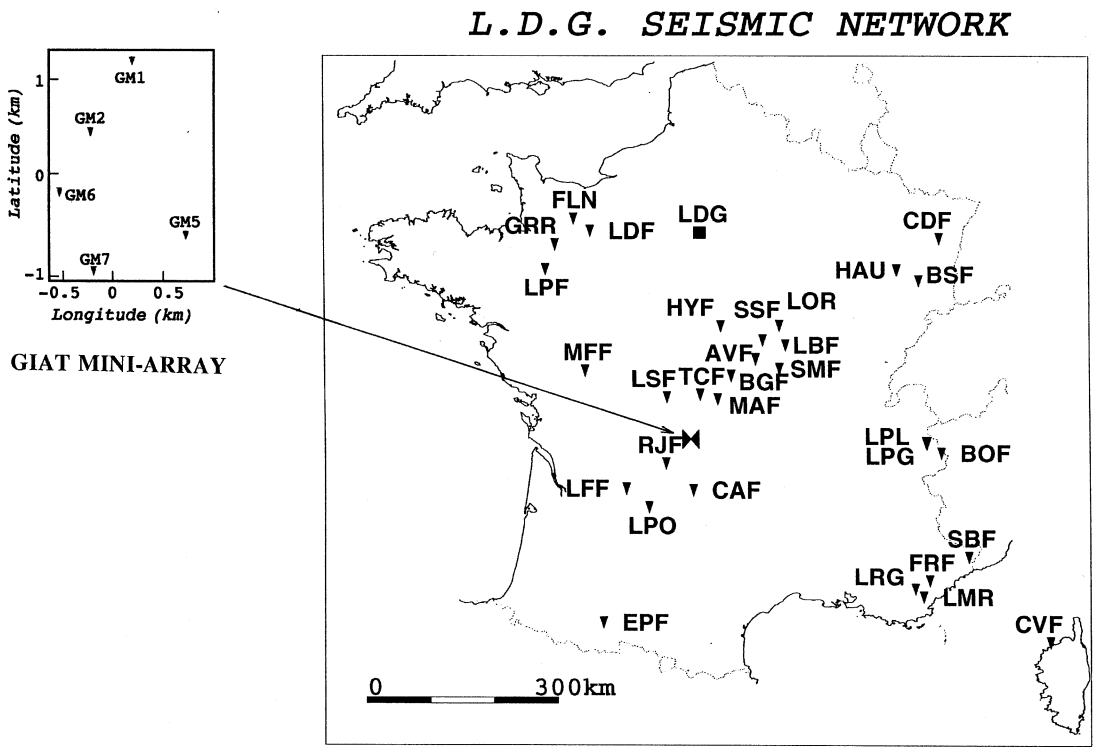


Fig. 1. Map showing the LDG national permanent network and the mini-array which is used in this study.

This mini-array recorded more than 100 teleseismic events and about 28 regional events during its 4-month operating period. This experience provided a useful data base for testing different automatic location methods (Cansi *et al.*, 1992). We showed that better results are obtained when we use the two-step correlation-method which computes first the arrival-time differences with a high precision (less than the sampling interval 0.02 s) and second the azimuth and propagation velocity corrected for statistically determined station anomalies.

2.1. *Data processing*

For regional events, the correlation functions which define the arrival-time differences often have several relative maxima with comparable amplitudes. This leads to an ambiguity in the arrival-time computation, which can be removed by testing the consistency of the time-delay set, using the consistency relationship:

$$\Delta t_{ij} + \Delta t_{jk} + \Delta t_{ki} = 0 \qquad (2.1)$$

Furthermore, this consistency can be used as a signal detector. When the studied time-window contains a seismic signal, the Root Mean Square of the residuals of the consistency relations is low (*i.e.*: less than the sampling interval: it is an estimate of the measurement accuracy). On the contrary, when it contains only noise, no consistency can be found

(*i.e.*: the RMS of consistency relations is high), because of the low signal correlation at the scale of the array.

Then, for each 4.5 s-time window and for different frequency-bands, we can estimate a probability of signal occurrence and, in the case of high probability, the corresponding velocity and azimuth. Some examples of these time-frequency functions are displayed in figs. 2 and 3. For each time window moving by a 1.5 s step and for each frequency-band, the velocity is shown (colour scale) only when the consistency is lower than 0.02 s.

We can see clearly that most of the seismic phases lead to consistent signals whose velocity is well defined for all the available frequency-bands. Nevertheless, some cases are more ambiguous:

– a phase cannot be precisely recognized because the velocity is not clearly defined (see *Sn*-phase in fig. 2);

– a false detection is obtained because part of the record contains consistent noise with a velocity compatible with the regional phase velocity (see noise in figs. 2 and 3 as an example).

Since these two kinds of problems cannot be easily modeled, we used a «learning system» approach based on a neural network to identify the different phases of each event without ambiguity.

2.2. *Phase identification*

For neural network applications, the classical programming efforts are transposed to the determination of the various authorised degrees of freedom described as follows (Lacoss *et al.*, 1991):

– *the data structure*: the first step is to extract from the data base the information that is strictly necessary for phase identification. Furthermore, those data have to be invariant by translation, rotation and dilatation, which precludes the analysis on a variable period. The solution we chose is thus to present as inputs and for each time sample the signal velocity for 5 frequency bands (from 0 to 12 Hz). The data whose corresponding consistency RMS is greater than an arbitrary threshold (*i.e.*: 0.02 s) are set to 0;

– *the network structure*: we only used multi-layered perceptrons, with a sigmoïd transfer function. They are indeed capable of building complex decision hyper-volumes in the hyper-space of the input data, thus realising an accurate classifier. Several tests led us to choose a 2 hidden-layer perceptron. The complexity of the system is due to the high non-linearity of the problem;

– *the learning function*: we chose as a learning function the «back propagation with momentum» method, which uses a gradient method to minimize the quadratic error between the expected and the observed results. Despite a long computing time, this allows a reliable and accurate learning convergence;

– *the example data base*: it was made of all the available events, except 3 of them on which the tests were made. In order to avoid incoherencies, the arrival time of each phase was picked on the time-frequency diagrams; a phase is thus declared present over the whole time-window between its arrival time and the arrival time of the following phase, in order to take into account the whole information of each phase, including its coda. For *Sg* phase, the window length is limited to 30 s. Each sample is presented 50 000 times in a random order;

– *the network topology*: the number of nodes in the hidden layers was determined empirically. Several networks were designed: in order to avoid over-training, we chose for each phase the simplest one which did not degrade the results. Finally, the *Pn* phase requires 24 + 9 units, the *Pg* phase 20 + 6, the *Sn* phase 16 + 6 and the *Sn* phase 20 + 6.

All the designing, learning and testing operations were performed using the neural simulator SNNS, developed by Stuttgart University

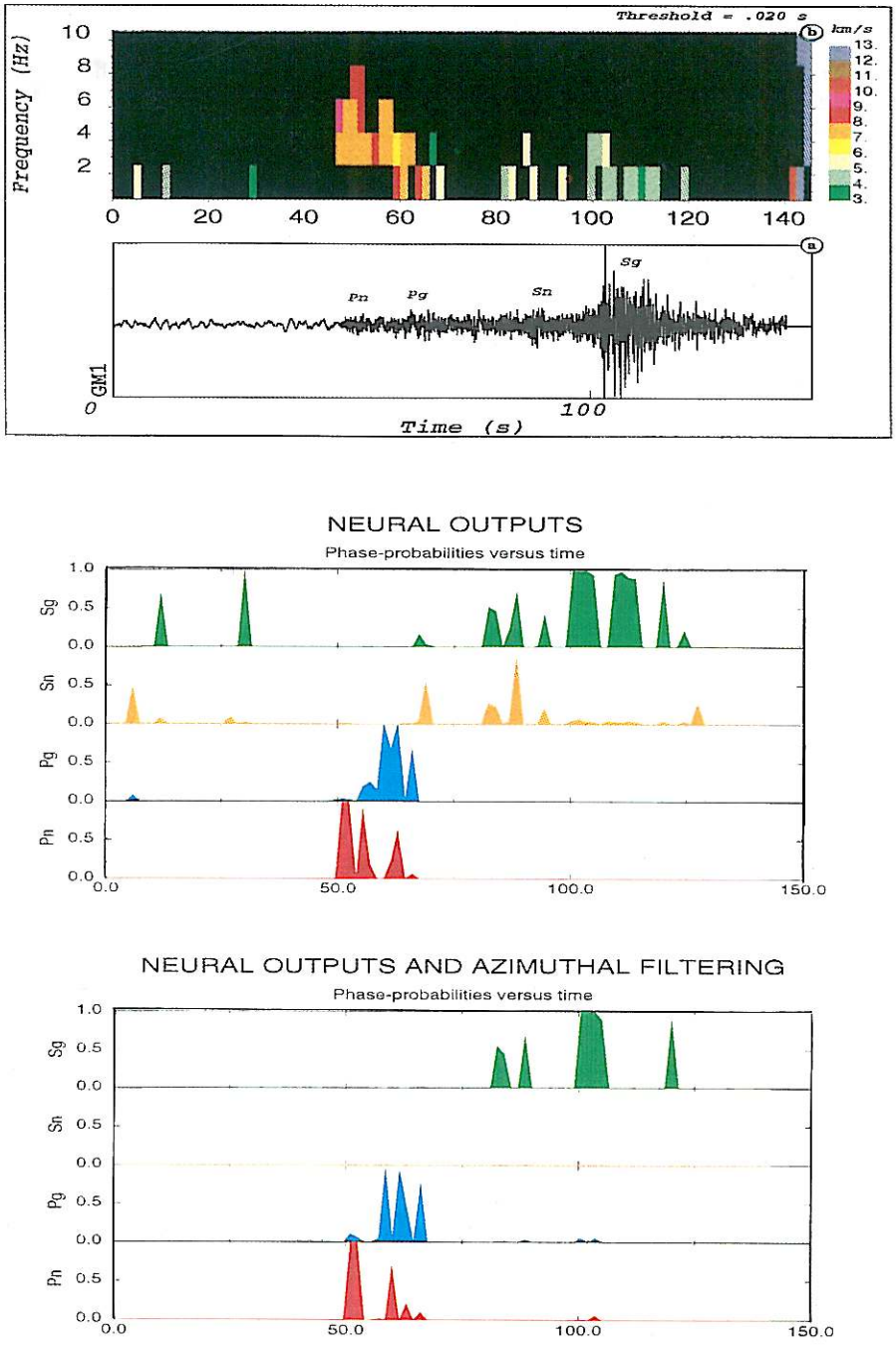
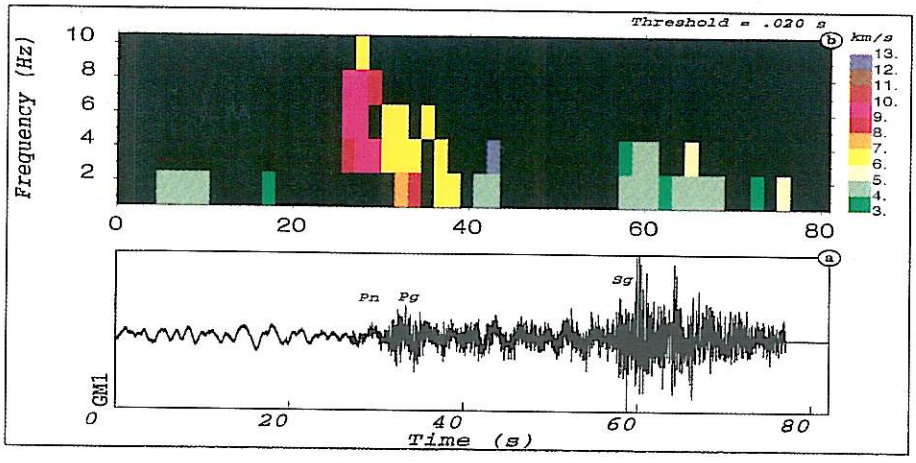


Fig. 2. Example of an event at the different processing levels: the time-frequency plot of the velocity (top), the 4 phase probablilty functions (middle) and the same after azimuth filtering (bottom).



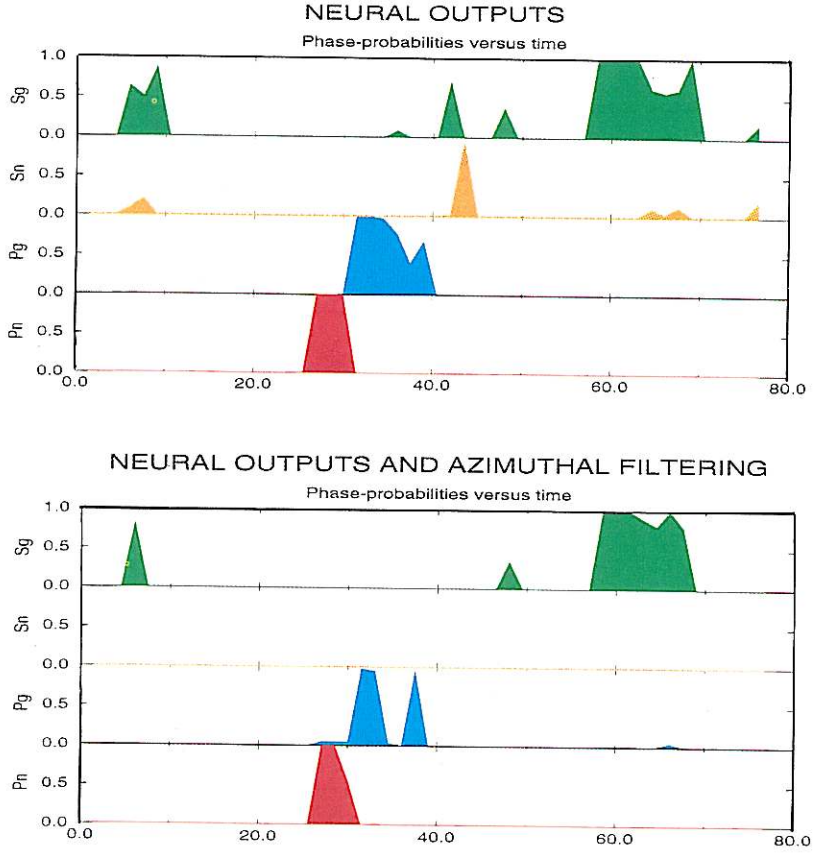


Fig. 3. See fig. 2.

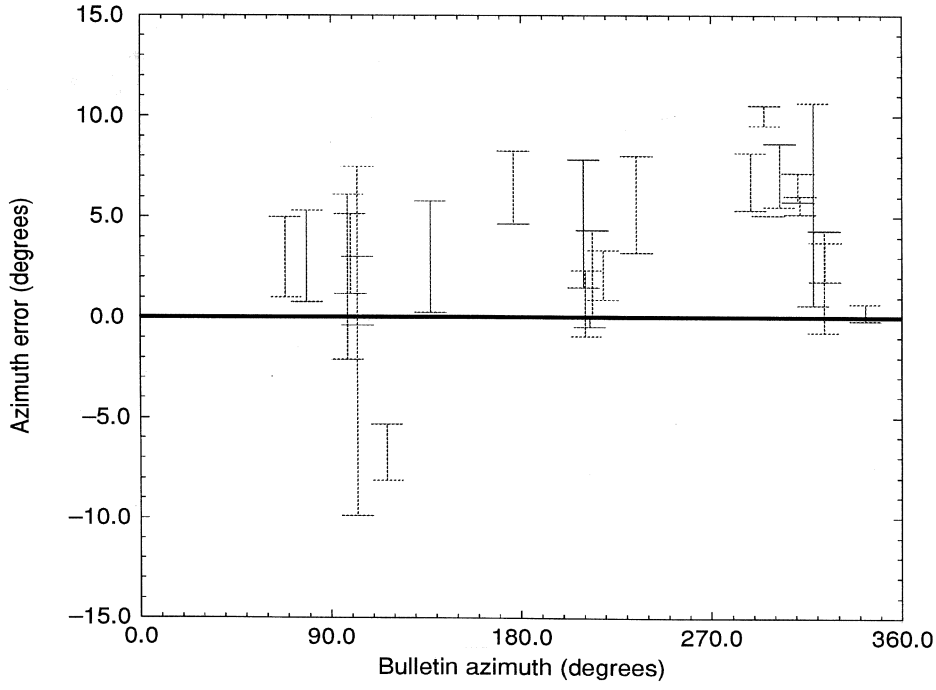


Fig. 4. Azimuth results: azimuth errors versus bulletins azimuths (references).

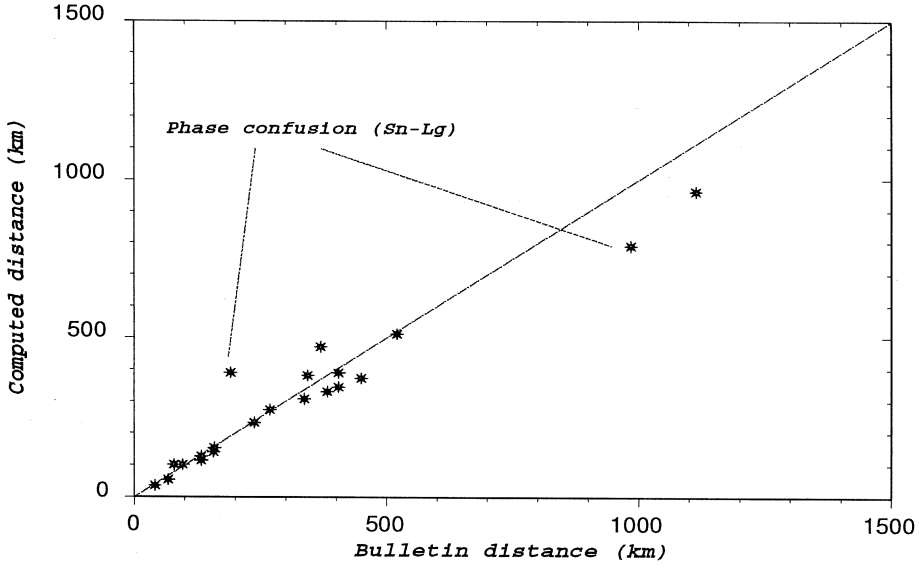


Fig. 5. Computed distance (this study) versus bulletin distance (reference).

(1992). The middle diagrams in figs. 2 and 3 show the neural outputs as a function of time for two events which were not in the learning data base.

2.3. *Distance estimation*

In order to remove the last false detections due to consistent noise, post data processing is needed to test the consistency of the results on the whole signal, by using the azimuthal information as described as follows:

– the first step is to compute an azimuth histogram with the possibility functions previously determined and to choose the most probable 20° wide interval. The average azimuth can then be calculated;
– the second step is to refine this approximation by determining the most probable 5° wide sub-interval for each phase. For each time sample, the probability of existence of a phase is set to 0 if the corresponding azimuth is out of those sub-intervals.

The bottom diagrams in figs. 2 and 3 show the phase characterization curves after this azimuthal filtering.

At this step, in the function describing the possibility of occurrence of each phase as a function of time, all the information which does not belong to the detected event has been removed. The last step – the distance evaluation – can now be performed.

Each function is differentiated to allow an easy detection of each phase by identifying the times where the derivative is greater than an arbitrary threshold (*i.e.*: 0.7). When only two phases are detected, the resulting distance is computed using the two times for which the product of the corresponding possibility functions is at its maximum. When more than two phases are detected, a least-square estimation of the distance is performed for each set of possible arrival-times. The best one is retained as the event distance.

2.4. *Results*

We have processed 28 regional events which occurred in Western Europe during the experiment. Their magnitudes ranged from $M_l = 2.2$ to $M_l = 5.3$ and their epicentral distances from 40 to 1200 km.

The azimuth error is shown as a function of the azimuth in fig. 4. It is clear that this method leads to a very accurate azimuth estimation: the mean value of the error bars is roughly 3 degrees.

For the 21 events for which a distance can be computed (*i.e.*: at least two phases have been recognized), the computed distance is shown versus the bulletin distance in fig. 5. Except two events for which confusions occurred between *Sn* and *Lg*-phases, a good agreement is obtained in the full distance range: the relative standard error is less than 20%. These results are illustrated in a regional map showing the location errors (fig. 6).

3. Conclusions and recommendations

We have tested a new method for automatic location of regional seismic events recorded on a mini-array. This method includes first a signal analysis process in order to evaluate the consistency of the relative time delays computed from the correlation functions, and second a neural network analysis to identify the different phases and compute the distance of the event.

The standard error of the azimuth estimation is less than 4 degrees for a set of events with distances ranging from 40 to 1200 km. Most of the *P*-phases are clearly identified by the system, but some confusion between *Sn* and *Lg* phases can be observed. When the cases of misidentification are removed, the method leads to an estimate of the distance with an uncertainty of about 20%.

We have shown that this method is able to produce an estimation of the location of regional events with a reasonable error. Further studies will be devoted to testing the method on a larger data set including the new design of the mini-array which is now composed of 10

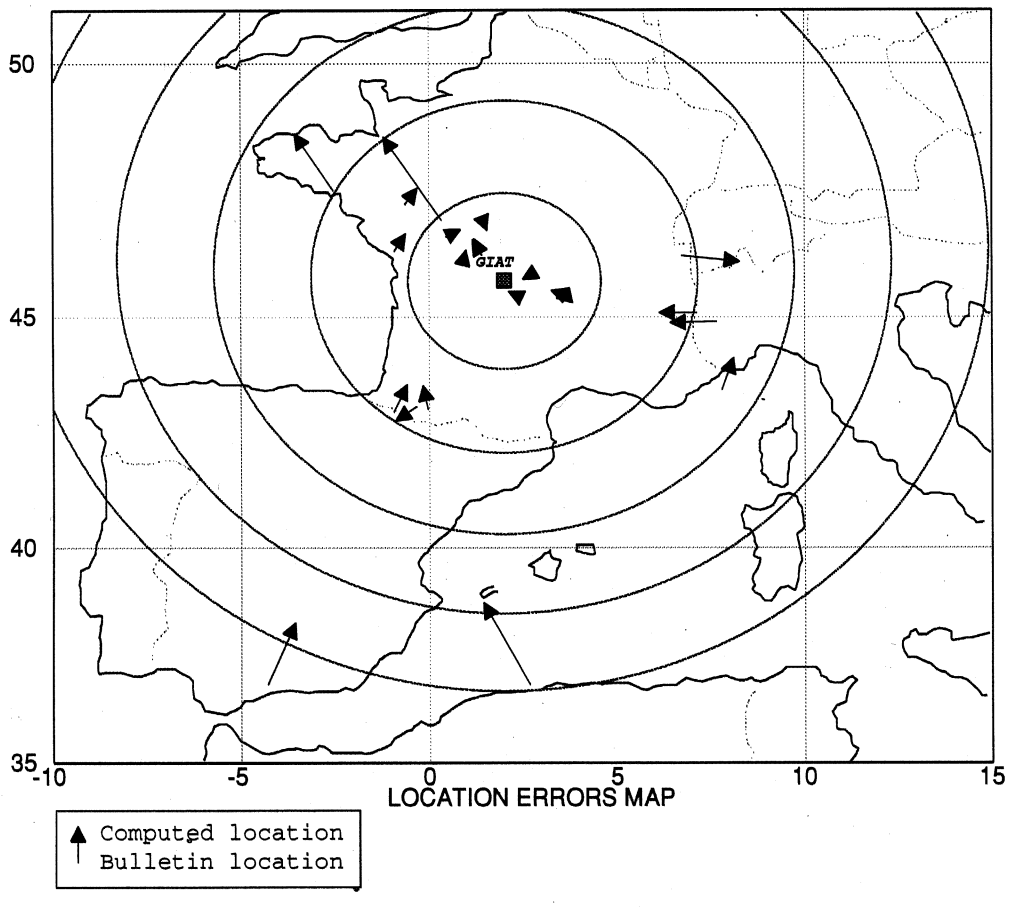


Fig. 6. Regional map showing the location errors for all the events.

stations, and for lower SNR ratios. Comparisons with azimuthal and velocity determinations computed from the new 3 components station and from the broad band *f-k* analysis will also be performed.

REFERENCES

CANSI, Y., J.L. PLANTET and J.P. MASSOT (1992): Epicentral azimuth determination by mini-array processing, *14th Annual DARPA Research Symposium*, 55-62.

HUSEBYE, E.S. (1969): Direct measurement of *dt/dΔ*, *BSSA*, **59** (2), 717-727.

LACOSS, R.T., S.R. CURTIS, R.K. CUNNINGHAM and M. SIBERT (1991): Seismic phase and event labeling using artificial neural network, *13th Annual DARPA Research Symposium*, 313-319.

MYKKELTVEIT, S., F. RINGDAL, T. KVÆRNA and R.W. ALEWINE (1990): Application of regional arrays in seismic verification, *Bull. Seismol. Soc. Am.*, **80**, 1777-1800.

UNIVERSITY STUTTGART (1992): *SNNS User Manual, Version 2.0*, ftp-anonymous: 129.69.21.11.

(received April 13, 1994;
accepted August 22, 1994)